

AD-A071 063

RICE UNIV HOUSTON TEX DEPT OF ELECTRICAL ENGINEERING
NONLINEAR SYSTEM IDENTIFICATION BASED ON A FOCK SPACE FRAMEWORK--ETC(U)
MAY 79 L V ZYLA, R J DE FIGUEIREDO
EE-TR-7901

F/G 9/4

N00014-79-C-0442

NL

UNCLASSIFIED

| OF |

AD
A071063



END
DATE
FILMED
8-79
DDC

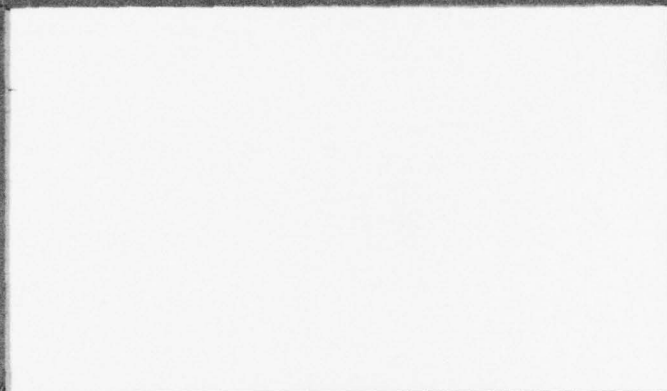
LEVEL II

42

RICE UNIVERSITY

GEORGE R. BROWN SCHOOL OF ENGINEERING

AD A C 7 1 0 6 3



DDC
RECEIVED
JUL 11 1979
C

DDC FILE COPY

DEPARTMENT OF ELECTRICAL ENGINEERING

HOUSTON, TEXAS

This document has been approved
for publication and sale; its
distribution is unlimited.

79

This document has been approved
for public release and sale; its
distribution is unlimited.

NONLINEAR SYSTEM IDENTIFICATION
BASED ON A FOCK SPACE FRAMEWORK

by

L.V. Zyla and Rui J.P. de Figueiredo

Rice University
EE Technical Report No. 7901

May, 1979



CONTENTS

	Page
ABSTRACT	1
1. Introduction	2
2. Developments Pertaining to Fock Spaces and Volterra Expansions	5
3. Hilbert Spaces of Nonlinear Operators	18
4. Nonparametric Nonlinear System Identification in the Noiseless Case: NonCausal Solution	20
4.1. Geometrical Considerations	21
4.2. Solution and Algorithm	24
5. Causal Solution	27
5.1. The Case of a Single Probing Input-Output Pair	30
5.2. The Case of Several Probing Input-Output Pairs	31
5.2.1. Strictly Causal Solution	31
5.2.2. ϵ -Causal Solution	36
6. Nonlinear System Identification in the Noisy Case	38
7. Computer Simulation Example	41
8. Conclusion	44
REFERENCES	45

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Rice Univ Tech. Report EE-7901	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NONLINEAR SYSTEM IDENTIFICATION BASED ON A FOCK SPACE FRAMEWORK.	5. TYPE OF REPORT & PERIOD COVERED Final report of one phase of a continuing problem	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) L.V. Zyla ■ Rui J.P./de Figueiredo	8. CONTRACT OR GRANT NUMBER(s) N000 14-79-C-0442 <i>mer</i>	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of Electrical Engineering Rice University, P.O. Box 1892 Houston TX 77001	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Task No. NR 042419	
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics & Probability Program (Code 436) Office of Naval Research Arlington VA 22217	12. REPORT DATE May 1979	
13. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) <i>(12) 48 p.</i>	14. NUMBER OF PAGES 46	
	15. SECURITY CLASS. (of this report) Unclassified	
	16. DECLASSIFICATION/DOWNGRADING SCHEDULE	
17. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED. <i>(14) EE-TR-7901</i>		
18. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
19. SUPPLEMENTARY NOTES This work was also supported in part by the NSF Grant ENG 74-17955 and by the RADC (Air Force Systems Command) Contract F 30602-78-C-0148.		
20. KEY WORDS (Continue on reverse side if necessary and identify by block number) Nonlinear systems; system identification; remote sensing; estimation theory; optimization; approximation theory; Fock spaces; Volterra functionals; operator theory; functional analysis.		
21. ABSTRACT (Continue on reverse side if necessary and identify by block number) <i>yields</i> A method is presented for the identification of a nonlinear system repre- sented by an operator $V:E \rightarrow Y$, where the input space E is a separable Hilbert space over the field of complex numbers and the output space Y is the Sobolev space $H_n^2(I)$ of complex-valued functions y on an interval I of the real line such that $d^k y/dt^k$, $k=0, \dots, n-1$, are absolutely continuous and $d^n y/dt^n \in L^2(I)$. <i>Hilbert-Squared</i> (Cont.)		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LE-014-6601

403 244 *DM*
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)
An element of L-Squared

20. ABSTRACT (CONTD.)

Our identification scheme is based on a given set of input-output pairs $\{(u_i, y_i): u_i \in E, y_i \in Y, i=1, \dots, m\}$ and on an appropriate description of the class C to which V is assumed to belong, a description which permits us to incorporate in the formulation of C our a-priori knowledge of the properties of V .

Let $y(t) = (Vu)(t)$ denote the output at time t corresponding to an input u . We express this in the form $y(t) = V_t(u)$, where we assume that V belongs to the Fock space $F_\rho(E)$ of order ρ over E . If $E = L_2(I)$, V can be represented by a Volterra functional expansion. We show and use the property that $F_\rho(E)$ is a reproducing kernel Hilbert space. Also, we construct a Hilbert space $B_n^2(I, F_\rho(E))$ of nonlinear operators from I to $F_\rho(E)$ and characterize the operator class C to which V belongs as an ellipsoidal class in $B_n^2(I, F_\rho(E))$.

The above developments permit us to obtain the solution to our nonlinear system identification problem as the solution to an appropriate minimum norm problem in $B_n^2(I, F_\rho(E))$. Procedures for obtaining both the noncausal and causal solutions are given. We also introduce the concept of " ϵ -causality", which is weaker than that of causality, and derive an ϵ -causal solution to our problem. The case when measurement errors are present is finally considered.

The results are illustrated by the computer simulation of a simple example in which very good agreement with the theory is obtained over a wide time-interval.

B.m. squared

epsilon

1. Introduction

Let S denote a system represented by a nonlinear* operator $V: E \rightarrow Y$, where E and Y are appropriate Hilbert spaces called respectively the "input" and "output" spaces. From now on, the norm and inner product in a given space, say H , will be denoted by $(\cdot, \cdot)_H$ and $\|\cdot\|_H$, the subscript being omitted when the space referred to is clear from the context.

In the present paper, we consider the problem of identifying S and hence V , based on a given set of input-output pairs $\{(u_i, y_i) | u_i \in E, y_i \in Y, i = 1, \dots, m\}$ (called "probing input-output pairs") and on the class C to which V is assumed to belong.

Our framework is such that the operator class C is defined by assuming a "finite gain" property for the operator V and taking into account the smoothness properties of the output. Furthermore, members of C are not required to be known up to a finite set of parameters. For this reason, we regard the system identification approach developed here as being a "non-parametric approach."

Specifically, we formulate the nonlinear system identification problem in its fullest generality as follows:

For a given $\tilde{V} \in C$, let $e_C(\tilde{V})$ denote a norm on the operator class C which measures an appropriate error in the approximation of \tilde{V} by members of the entire class C . Then, under no measurement noise conditions, \hat{V} is our best estimate of the operator V (which is to be identified) if it is the solution of the following problem:

* By "nonlinear" we mean "not necessarily linear."

Accession For	
NTIS GAO&I	<input checked="" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

Problem 1. Given $u_i \in E$, $y_i \in Y$, $i = 1, \dots, m$ and an appropriate operator class C , find \hat{V} as the solution of

$$\min\{e_C(\tilde{V})\} \quad (1.1)$$

$$\tilde{V} \in C$$

$$\tilde{V}(u_i) = y_i, \quad i = 1, \dots, m. \quad \parallel^*$$

If the output measurements are corrupted by noise, we model the relation between the probing input-output pairs (u_i, y_i) , $i = 1, \dots, m$, by

$$y_i = V(u_i) + v_i, \quad i = 1, \dots, m, \quad (1.2)$$

where $v_i \in Y$ satisfy**

$$(v_i, v_j)_Y = q_i \delta_{ij}, \quad i, j = 1, \dots, m, \quad (1.3)$$

q_i being positive constants, and δ_{ij} = Kronecker delta. The set of equations (1.3) may be interpreted as there being no correlation between the noise present in the i th and j th measurements. In this case our best estimate \hat{V} of V is the solution of:

Problem 2. Minimize

$$J(V, \sum_{i=1}^m q_i^{-1} \|V(u_i) - y_i\|_Y^2). \quad (1.4)$$

where J is a criterion which optimizes the estimate with respect to the class properties of V as well as noise. \parallel

In what follows, we assume that the input space E is a separable Hilbert space over the field of complex numbers; Y is the Sobolev space $H_n^2(I)$ of complex-valued functions g on an interval I of the real line such that

* Ends of formal statements will be signified by the symbol \parallel .

** If the noise belongs to some bigger space than Y , then we define v_i as the projection of the noise on Y .

$g^{(k)} (\triangleq d^k g / dt^k)$, $k = 0, 1, \dots, n-1$, are absolutely continuous and $g^{(n)} \in L^2(I)$; and V belongs to a Hilbert space of nonlinear operators $B_n^2(I; F_\rho(E))$, which we call a Bochner-Sobolev space of order n over $F_\rho(E)$, where $F_\rho(E)$ is a Fock space of order ρ over E .

In sections 2 and 3, we define the spaces $F_\rho(E)$ and $B_n^2(I, F_\rho(E))$, and present mathematical developments pertaining to them needed in later sections. We show that $F_\rho(E)$ is a reproducing kernel Hilbert space and we establish useful links between $F_\rho(E)$ and $B_n^2(I, F_\rho(E))$ on one hand, and their particular manifestations in the form of spaces of Volterra functionals and of Volterra operators on the other hand.

Section 4 is devoted to a detailed formulation and solution of Problem 1 for the general case of a noncausal operator. In Section 5, these results are particularized to the causal case and (after the introduction of the concept of "ε-causality") to the ε-causal case. Problem 2 is solved in Section 6 and an example presented in Section 7.

The present approach to the system identification problem is similar to the one proposed by de Figueiredo and Caprihan [1], [2] for the identification of linear systems, with the basic difference that in the linear case, the space, to which the operator to be identified belonged, was assumed to be the space of "trace class" operators.

Referring to contributions of other authors relevant to the developments in the present paper, special mention should be made of the early work of A.V. Balakrishnan [3] and others (see references in [2]) on the identification of nonlinear systems from input-output data, as well as the recent contributions of F.J. Beutler and W.L. Root (see [4] and references therein) on the identification of linear as well as polynomial systems much along the lines of the present paper. Other papers on related topics are the ones by J.L. Franklin [5], E. Mosca [6], and W. Porter [7] [8] [9].

We now proceed to construct the mathematical framework used in the solution of Problems 1 and 2.

2. Developments Pertaining to Fock Spaces and Volterra Expansions

Let E be the separable Hilbert space introduced previously, and denote by $B_n(E; \mathbb{C})$ (where \mathbb{C} denotes the space of complex numbers) the Banach space of n -linear symmetric bounded forms $A : E \times E \times \dots \times E \rightarrow \mathbb{C}$ where

$$|A(u_1, \dots, u_n)| \leq M \|u_1\| \dots \|u_n\| \quad (2.1)$$

for some $M \in \mathbb{R}_+^1$. For any $A \in B_n(E; \mathbb{C})$, define a new functional $P_A : E \rightarrow \mathbb{C}$ by

$$P_A(u) = A(u, \dots, u) \quad (2.2)$$

and a norm on P_A by

$$\|P_A\| = \sup_{\|u\|=1} |A(u, \dots, u)| \quad (2.3)$$

The space of functionals $\{P_A | A \in B_n(E; \mathbb{C})\}$ is a normed linear space, and we denote by $P(^n E)$ the Banach space of continuous n -homogeneous polynomials obtained by completing $\{P_A | A \in B_n(E; \mathbb{C})\}$ in the (2.3) norm.

For $A \in B_n(E; \mathbb{C})$ define the norms $\|\cdot\|_H$ and $\|\cdot\|_B$ as follows (the subscripts H and B stand for Hilbert and Banach respectively and do not refer to specific spaces)

$$\|A\|_H = \sup_{\|u_1\|=1, \dots, \|u_n\|=1} |A(u_1, \dots, u_n)|, \quad (2.4)$$

$$\|A\|_B = \sup_{\|u\|=1} |A(u, \dots, u)|, \quad (2.5)$$

and it is immediately obvious that $\|A\|_B \leq \|A\|_H$.

It is not difficult to conclude that for symmetric n -linear forms one can bound $\|A\|_H \leq c_n \|A\|_B$ for some constant c_n dependent only on n [10]. So $\|\cdot\|_H$ and $\|\cdot\|_B$ are equivalent norms on the space $B_n(E; \phi)$.

Let H^n denote the tensor product of a Hilbert space H with itself n times [11, 12] and let $e_1 \otimes e_2 \otimes \dots \otimes e_n$ be a decomposable element of E^n where $e_i \in E$. Then from the definition of E^n it follows that $\|e_1 \otimes e_2 \otimes \dots \otimes e_n\|_{E^n} = \|e_1\| \dots \|e_n\|$ and $(e_1 \otimes e_2 \otimes \dots \otimes e_n) [u_1, \dots, u_n] = (e_1, u_1)_E \dots (e_n, u_n)_E$ for $e_1, \dots, e_n, u_1, \dots, u_n \in E$. Here $[\cdot, \dots, \cdot]$ is a member of the n -th order Cartesian product of E . Introduce i : $e_1 \otimes e_2 \otimes \dots \otimes e_n \rightarrow (e_1, \cdot)_E \dots (e_n, \cdot)_E$. Then i maps the decomposable elements of E^n into $P(^nE)$. It is clear that

$$\|i(e_1 \otimes \dots \otimes e_n)\| = \|(e_1, \cdot)_E \dots (e_n, \cdot)_E\|_B \leq \|e_1\| \dots \|e_n\|. \quad (2.6)$$

Define i by linearity on all finite linear combinations of decomposable elements of E^n and, using equation (2.6), it is easy to demonstrate [12, 13] that i is a bounded linear map of finite linear combinations of decomposable elements of E^n into $P(^nE)$. Finally, by continuity extend i to all of E^n . Define a new map $S_n: E^n \rightarrow E^n$ by

$$S_n = \frac{1}{n!} \sum_{\bar{\sigma} \in P_n} \bar{\sigma}, \quad (2.7)$$

where P_n is the permutation group of n integers and $\bar{\sigma}$ acts on decomposable elements of E^n by

$$\bar{\sigma} (e_1 \otimes e_2 \otimes \dots \otimes e_n) = e_{\sigma_1} \otimes e_{\sigma_2} \otimes \dots \otimes e_{\sigma_n}. \quad (2.8)$$

Since E is separable, there is $\{\varphi_k\}_{k=1}^{\infty}$, an orthonormal basis for E . Then it is well known [12] that $\{\varphi_{i_1} \otimes \varphi_{i_2} \otimes \dots \otimes \varphi_{i_n}\}_{i_1=1, \dots, i_n=1}^{\infty, \dots, \infty}$ is an orthonormal basis for E^n . Clearly,

$$\bar{\sigma} \text{ may be defined on all of } E^n \text{ by } \bar{\sigma} \left(\sum_{i_1=1}^{\infty} \dots \sum_{i_n=1}^{\infty} c_{i_1, \dots, i_n} \varphi_{i_1} \otimes \dots \otimes \varphi_{i_n} \right) = \sum_{i_1=1}^{\infty} \dots \sum_{i_n=1}^{\infty} c_{i_1, \dots, i_n} \varphi_{i_{\sigma_1}} \otimes \dots \otimes \varphi_{i_{\sigma_n}}, \quad (2.8); \text{ and furthermore, one can}$$

show that $\bar{\sigma}$ is a bounded linear operator on E^n into E^n . It then follows [12] that S_n defined in equation (2.7) is a projection operator in E^n , i.e. $S_n^2 = S_n$ and $S_n^\dagger = S_n$ (where the superscript \dagger denotes the adjoint).

The subspace $S_n E^n$ of E^n is called the symmetric tensor product of E of order n . For the case of $E = L_2(\mathbb{R}^1)$, E^n is just $L_2(\mathbb{R}^1) \otimes \dots \otimes L_2(\mathbb{R}^1)$ and it can be shown that there exists an isomorphism J between E^n and $L_2(\mathbb{R}^n)$

such that J sends the element $\sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} e_{i_1} \otimes \dots \otimes e_{i_n}$ into

$h(t_1, \dots, t_n)$, where $h \in L_2(\mathbb{R}^n)$ and $\{e_i\}_{i=1}^{\infty}$ is an orthonormal basis for

$L_2(\mathbb{R}^1)$. $S_n E^n$ is then just the subspace of $L_2(\mathbb{R}^n)$ of functions left invariant under any permutation of their variables. Furthermore, one can show [12] that:

$$\sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} e'_{i_1}(u_1) \dots e'_{i_n}(u_n) = \int_{\mathbb{R}^1}^{(n)} h(t_1, \dots, t_n) u_1(t_1) \dots u_n(t_n) dt_1, \dots, dt_n, \quad (2.9)$$

* The superscript * denotes complex conjugate.

where

$$e'_i(u) = \int_R e_i^*(t) u(t) dt, \quad (2.10)$$

where $u_1, u_2, \dots, u_n \in L_2(R)$. To complete our construction we look at the space $i(S_n E^n) \subset P({}^n E)$. Certainly $i(S_n E^n)$ is a linear space and since i is 1-1 on $S_n E^n$ [13] we can define an inner product between any two elements $u, v \in i(S_n E^n)$ by

$$\langle u, v \rangle = (i^{-1}(u), i^{-1}(v))_{E^n} \quad (2.11)$$

Now we invoke the following theorem:

Theorem 2.1

If i is a 1-1 bounded linear operator from a Hilbert space H into a Banach space B then the image $i(H) \subset B$ is a Hilbert space with the inner product

$$\langle u, v \rangle = (i^{-1}(u), i^{-1}(v))_H \quad \parallel \quad (2.12)$$

Proof: It is clear that $i(H)$ is a vector space in B and furthermore it is an inner product space with inner product (2.12). To show that $i(H)$ is actually a Hilbert space we must demonstrate completeness. Let $\{b_i\}$ be a Cauchy sequence in $i(H)$, ie $\forall \epsilon > 0 \exists N(\epsilon)$ such that $\langle b_i - b_j, b_i - b_j \rangle < \epsilon$ for $i, j \geq N_\epsilon$. But

$$\begin{aligned} \langle b_i - b_j, b_i - b_j \rangle &= (i^{-1}(b_i - b_j), i^{-1}(b_i - b_j))_H = \\ &= (i^{-1}(b_i) - i^{-1}(b_j), i^{-1}(b_i) - i^{-1}(b_j))_H \end{aligned} \quad (2.13)$$

so that $\{i^{-1}(b_i)\}$ is a Cauchy sequence in H . But H is complete so there exists $h \in H$ such that $i^{-1}(b_i) \rightarrow h$. Therefore

$$\langle b_{i^{-1}(h)}, b_{i^{-1}(h)} \rangle = (i^{-1}(b_{i^{-1}(h)}), i^{-1}(b_{i^{-1}(h)}))_H = \\ (i^{-1}(b_i) - h, i^{-1}(b_i) - h) \rightarrow 0. \quad \parallel \quad (2.14)$$

We can now say that $i(S_n E^n)$ is a Hilbert space of n -homogeneous symmetric polynomials on E . We denote $i(SE^n)$ by $P_H^{(n)E}$ and $P_H^{(n)E}$ is known in the literature as the Hilbert-Schmidt polynomials on E [13]. From now on we will denote the inner product in $P_H^{(n)E}$ by $(\cdot, \cdot)_{E_n}$. We state the following propositions connecting tensor products of L_2 spaces and the Hilbert-Schmidt polynomials.

Proposition 2.1

Let M be a measure space and $E = L_2(M)$, then $P \in P_H^{(n)E}$ if and only if there exists $h_n \in L_2(M \times \dots \times M)$ such that

$$P(u) = \int_{M \times \dots \times M} \dots \int_{M \times \dots \times M} h_n(t_1, \dots, t_n) u(t_1) \dots u(t_n) d\mu_n, \quad (2.15)$$

h_n is symmetric and

$$\|P\|_{E_n} = \|h_n\|_{L_2(M \times \dots \times M)}. \quad \parallel \quad (2.16)$$

Proof: See Reed and Simon [12]. \parallel

Proposition 2.2

Given an orthonormal basis $\{e_i\}_{i=1}^\infty$ of E , each $P \in P_H^{(n)E}$ is uniquely expressed as a limit in $\|\cdot\|_{E_n}$ norm by

$$P = \sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} e'_{i_1} \dots e'_{i_n} \quad (2.17)$$

with symmetric coefficients $c_{i_1, \dots, i_n} \in \mathbb{C}$ and

$$\|P\|_{E_n}^2 = \sum_{i_1, \dots, i_n}^{10} |c_{i_1, \dots, i_n}|^2. \quad (2.18)$$

Here, $e'_i(u) = (e_i, u)_E$ for $u \in E$. If $R \in P_H^n(E)$ and

$$R = \sum_{i_1, \dots, i_n} d_{i_1, \dots, i_n} e'_{i_1}, \dots, e'_{i_n} \quad (2.19)$$

then

$$(P, R)_{E_n} = \sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n}^* d_{i_1, \dots, i_n}. \quad (2.20)$$

Proof: See Dwyer [13]. \parallel

We will now define the Fock Space [12, 13, 14] of functionals which will be the mathematical framework within which Volterra expansions will be considered.

Definition 2.1 Let ρ be some fixed real number > 0 . The Fock space of order ρ , denoted $F_\rho(E)$, is the space of sequences (f_0, f_1, f_2, \dots) where $f_0 \in \mathcal{C}$ and $f_n \in P_H^n(E)$, such that

$$\sum_{n=0}^{\infty} \frac{\rho^n}{n!} \|f_n\|_{E_n}^2 < \infty. \quad (2.21)$$

It is well known [12] from the construction of Cartesian products of Hilbert spaces that F_ρ is a Hilbert space. We will denote the inner product between two elements $f = (f_0, f_1, \dots)$ and $g = (g_0, g_1, \dots)$ of $F_\rho(E)$

$$\text{by} \quad \langle f, g \rangle_{F_\rho} = \sum_{n=0}^{\infty} \frac{\rho^n}{n!} (f_n, g_n)_{E_n} \quad (2.22)$$

What we now will show is that $F_\rho(E)$ can be considered to be a space of functionals on E by the relation*

$$f(u) = \sum_{n=0}^{\infty} \frac{f_n(u)}{n!} \quad (2.23)$$

* Thus an element $f \in F_\rho(E)$ will be viewed both as a sequence such as in the Definition 2.1 and a functional on E the evaluation of which is defined by (2.23).

for $f \in F_p(E)$. Suppose that we have an element $f_2 \in P_H(^2E)$; then we know by Proposition 2.2 that

$$f_2 = \sum_{i_1, i_2} c_{i_1, i_2} e'_{i_1}(\cdot) e'_{i_2}(\cdot) \quad (2.24)$$

and

$$\|f_2\|_{E_2} = \sum_{i_1, i_2} |c_{i_1, i_2}|^2. \quad (2.25)$$

But f_2 is also an element of $P(^2E)$ so we can consider the Banach norm (equation 2.5) of f_2 . By definition, $\|f_2\|_B = \sup_{\|u\|_E=1} |f_2(u)|$ and by (2.24),

$$\|f_2\|_B = \sup_{\|u\|_E=1} \left| \sum_{i_1, i_2} c_{i_1, i_2} e'_{i_1}(u) e'_{i_2}(u) \right|. \quad (2.26)$$

Since $\{e_i\}_{i=1}^\infty$ is a basis for E we can write $u = \sum_{i=1}^\infty \alpha_i e_i$, (2.27)

where

$$\alpha_i = (u, e_i)_E = e'_i(u).$$

Applying the Cauchy-Schwartz inequality to the right-hand side of (2.26), we get:

$$\begin{aligned} \left| \sum_{i_1, i_2} c_{i_1, i_2} e'_{i_1}(u) e'_{i_2}(u) \right| &\leq \sum_{i_1, i_2} |c_{i_1, i_2}| |e'_{i_1}(u)| |e'_{i_2}(u)| \leq \\ &\left(\sum_{i_1, i_2} |c_{i_1, i_2}|^2 \right)^{1/2} \left(\sum_{i_1, i_2} |e'_{i_1}(u)|^2 |e'_{i_2}(u)|^2 \right)^{1/2} = \end{aligned}$$

$$\left(\sum_{i_1, i_2} |c_{i_1, i_2}|^2 \right)^{1/2} \left(\sum_{i_1=1}^{\infty} |\alpha_{i_1}|^2 \right)^{1/2} \cdot \left(\sum_{i_2=1}^{\infty} |\alpha_{i_2}|^2 \right)^{1/2}, \quad (2.28)$$

and since $\|u\| = 1$ in (2.26), this implies $\sum_{i=1}^{\infty} |\alpha_i|^2 = 1$. Using this last

equation in (2.28) we finally obtain

$$\sup_{\|u\|=1} \left| \sum_{i_1, i_2} c_{i_1, i_2} e'_{i_1}(u) e'_{i_2}(u) \right| \leq \left(\sum_{i_1, i_2} |c_{i_1, i_2}|^2 \right)^{1/2} = \|f_2\|_{E_2}. \quad (2.29)$$

Stating these results in a proposition for arbitrary n , we get

Proposition 2.3

$$\text{If } f_n \in P_H({}^n E) \text{ then } \|f_n\|_B \leq \|f\|_{E_n} \quad (2.30)$$

$$\text{where } \|f_n\|_B = \sup_{\|u\|_E=1} |f_n(u)|.$$

We are now in position to show that relation (2.23) defines a bounded functional on E . If $f_n \in P_H({}^n E)$, then

$$\begin{aligned} |f_n(u)| &= \|u\|^n |f_n(\frac{u}{\|u\|})| \leq \|u\|^n \sup_{\|v\|=1} |f_n(v)| \\ &\leq \|u\|^n \|f_n\|_B \leq \|u\|^n \|f_n\|_{E_n}. \end{aligned} \quad (2.31)$$

We can see from (2.23), (2.30) and (2.31) that

$$\begin{aligned} |f(u)| &= \left| \sum_{n=0}^{\infty} \frac{f_n(u)}{n!} \right| \leq \sum_{n=0}^{\infty} \left| \frac{f_n(u)}{n!} \right| \leq \sum_{n=0}^{\infty} \sqrt{\frac{1}{\rho^n}} \frac{\|u\|^n}{n!} \sqrt{\rho^n} \|f_n\|_{E_n} \leq \\ &\left(\sum_{n=0}^{\infty} \frac{\|u\|^{2n}}{\rho^n n!} \right)^{1/2} \left(\sum_{n=0}^{\infty} \frac{\|f_n\|_{E_n}^2}{n! \rho^n} \right)^{1/2} \leq e^{\frac{\|u\|^2}{2\rho}} \left(\sum_{n=0}^{\infty} \frac{\|f_n\|_{E_n}^2}{n! \rho^n} \right)^{1/2} \end{aligned} \quad (2.32)$$

But since we are dealing with $f \in F_\rho(E)$, we can write (2.32) as

$$|f(u)| \leq e^{\frac{\|u\|^2}{2\rho}} \cdot \|f\|_{F_\rho} < \infty. \quad (2.33)$$

This allows us to state the major theorem of this section.

Theorem 2.2

If $f \in F_\rho(E)$ then $f = \sum_{n=0}^{\infty} \frac{f_n}{n!}$ is an entire function of bounded type,

ie. f takes bounded sets into bounded sets. If $D^n f(u)$ denotes the n -th Frechet derivative of f at u then

$$D^n f(0) = f_n. \quad (2.34)$$

The class of functions $F_\rho(E)$ is a Hilbert space with the inner product

$$\langle \cdot, \cdot \rangle_{F_\rho} \text{ given by } \langle f, g \rangle_{F_\rho} = \sum_{n=0}^{\infty} \rho^n \frac{1}{n!} \langle D^n f(0), D^n g(0) \rangle_{E_n}. \quad \parallel \quad (2.35)$$

Proof: Equation (2.34) establishes the first statement and the proof of the remainder of the theorem may be found in Dwyer [13]. \parallel

The construction of the space $F_\rho(E)$ for an arbitrary Hilbert space E is due to Dwyer [13,14].

The construction of $F_\rho(E)$ when $E = L_2(M)$, where M is a measure space was well known [12] before Dwyer. The characterization of $F_\rho(E)$ when $E = L_2(R)$ is still a most useful and easily understood Fock space.

Before we conclude this section, let us introduce a map which will later prove to be exceedingly useful. If $u \in E$ then the n -th order polynomial $\overbrace{u'(\cdot) \dots u'(\cdot)}^{(n)}$ on E is certainly an element of $P_H^{(n)}(E)$. Define $\exp(u) \in F_\rho(E)$ to be the functional

$$f = \sum_{n=0}^{\infty} \frac{f_n}{n!}, \quad (2.36)$$

$$\text{where } f_n = \underbrace{u'(\cdot) \cdots u'(\cdot)}_n.$$

It is clear that for $v \in E$

$$(\exp(u))(v) = e^{(u,v)_E} \quad (2.37)$$

and that

$$\|\exp(u)\|_{F_p} = e^{(\rho \|u\|_E^2)/2} \quad (2.38)$$

Now if $f_n \in P_H^{(n)}(E)$, then we may write

$$f_n = \sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} e'_{i_1} \cdots e'_{i_n}. \quad (2.39)$$

Using equation (2.27) we get for $\forall u \in E$, $\underbrace{u'(\cdot) \cdots u'(\cdot)}_n =$

$$\sum_{i_1, \dots, i_n} \alpha_{i_1} \cdots \alpha_{i_n} e'_{i_1} \cdots e'_{i_n}. \quad (2.40)$$

Therefore by Proposition 2.2 we conclude that

$$(f_n, \underbrace{u'(\cdot) \cdots u'(\cdot)}_n)_{E_n} = \sum_{i_1, \dots, i_n} c_{i_1, \dots, i_n} \alpha_{i_1} \cdots \alpha_{i_n} = f_n(u). \quad (2.41)$$

We can now state the following important result.

Proposition 2.4

$$\text{If } f \in F_p(E) \text{ and } u \in E, \text{ then } f(u) = \langle f, \exp(\frac{u}{\rho}) \rangle_{F_p}. \quad (2.42)$$

Proof: By definition and equation (2.41),

$$f(u) = \sum_{n=0}^{\infty} \frac{f_n(u)}{n!} = \sum_{n=0}^{\infty} \frac{(f_n, u(\cdot) \dots u(\cdot))_{E_n}}{n!} =$$

$$\sum_{n=0}^{\infty} \frac{\rho^n}{n!} (f_n, \overbrace{u(\cdot) \dots u(\cdot)}^n)_{E_n} = \langle f, \exp(\frac{u}{\rho}) \rangle_{F_\rho} \quad \parallel \quad (2.42a)$$

It is worthwhile to rederive equation (2.33) in light of the above proposition. It is clear that

$$|f(u)| = \langle f, \exp(\frac{u}{\rho}) \rangle_{F_\rho} \leq \|f\|_{F_\rho} \|\exp(\frac{u}{\rho})\|_{F_\rho} \quad (2.43)$$

by (2.42) and the Cauchy-Schwartz inequality. By equation (2.38) we may rewrite (2.43) as

$$|f(u)| \leq \|f\|_{F_\rho} e^{\|u\|^2/2\rho} \quad (2.44)$$

which is just inequality (2.32). We have actually shown more in Proposition 2.4 than we set out to do. We have proven that point evaluation is a continuous functional on the space $F_\rho(E)$. Equation (2.42a) identifies the representer of the point evaluation at u as the element $\exp(\frac{u}{\rho})$. For future reference we state this observation in the form of a theorem.

Theorem 2.3

The Fock space $F_\rho(E)$ is a Reproducing Kernel Hilbert space with the

reproducing kernel $e^{(u,v)_{E/\rho}}$ for $u, v \in E$. ||

Suppose that our Hilbert space $E = L_2(R)$. What would the Fock Space of $L_2(R)$ look like? Proposition 2.1 tells us that $f_n \in P_H^n(L_2(R))$ has the form

$$f_n(u) = \int_{R^n} \frac{(n)}{\cdot} h_n(t_1, \dots, t_n) u(t_1) \dots u(t_n) dt_1 \dots dt_n. \quad (2.45)$$

Thus since $f(u) = \sum_{n=0}^{\infty} \frac{f_n(u)}{n!}$, we may write

$$f(u) = h_0 + \frac{1}{1!} \int_R h_1(t_1) u(t_1) dt_1 + \frac{1}{2!} \iint_{R^2} h_2(t_1, t_2) u(t_1) u(t_2) dt_1 dt_2 + \\ \frac{1}{3!} \iiint_{R^3} h_3(t_1, t_2, t_3) u(t_1) u(t_2) u(t_3) dt_1 dt_2 dt_3 + \dots, \quad (2.46)$$

where $h_0 \in \mathbb{C}$, $h_1 \in L_2(R)$, $h_2 \in L_2(R^2)$, etc.

By definition,

$$\|f\|_{F_p}^2 = \sum_{n=0}^{\infty} \frac{\rho^n}{n!} \int_{R^n} \frac{(n)}{\cdot} |h_n(t_1, \dots, t_n)|^2 dt_1 \dots dt_n. \quad (2.47)$$

A question which might naturally arise is how much error we incur

by truncating a Volterra series. In other words, if $f(u) = \sum_{n=0}^{\infty} \frac{f_n(u)}{n!}$

and we approximate f by $\sum_{n=0}^N \frac{f_n(u)}{n!}$ how can we bound the error of such

an approximation in a pointwise sense? We know that

$$\left| f(u) - \sum_{n=0}^N \frac{f_n(u)}{n!} \right| = \left| \sum_{n=N+1}^{\infty} \frac{f_n(u)}{n!} \right| \leq \sum_{n=N+1}^{\infty} \|u\|^n \frac{\|f_n\|}{n!} E_n \\ \left(\sum_{n=N+1}^{\infty} \frac{\|u\|^{2n}}{\rho^n n!} \right)^{1/2} \left(\sum_{n=N+1}^{\infty} \frac{\rho^n}{n!} \|f_n\|^2 E_n \right)^{1/2} \leq \|f\|_{F_p} \left(\sum_{n=N+1}^{\infty} \frac{\|u\|^{2n}}{\rho^n n!} \right)^{1/2}. \quad (2.48)$$

So if we define the truncated exponential function

$$e_N(z) = \sum_{k=N}^{\infty} \frac{z^k}{k!}, \quad (2.49)$$

we can say that the error satisfies

$$\left| f(u) - \sum_{n=0}^N \frac{f_n(u)}{n!} \right| \leq \|f\|_{F_p} \sqrt{e_{N+1}(\|u\|^2/p)}. \quad (2.50)$$

The goodness of the truncated approximation is preserved only in a small neighborhood about the origin and deteriorates for large values of $\|u\|_E$.

3. Hilbert Spaces of Nonlinear Operators

We now proceed to introduce spaces of Hilbert-valued functions on an interval I of the real line.

Definition 3.1 $B_O^2(I, F_\rho(E))$ is the space of functions (operators) V from I to $F_\rho(E)$, where V is strongly measurable with respect to ordinary Lebesgue measure μ on I and satisfies

$$\int_I \|V(t)\|_{F_\rho(E)}^2 d\mu < \infty. \quad (3.1)$$

We will often denote the value $V(t) \in F_\rho(E)$ by V_t , and also replace $d\mu$ by dt with the understanding that all integrations are in the senses of Lebesgue or Bochner.

Note that if $u \in E$ and y is the function on I defined by

$$y(t) = V_t(u), \quad (3.2)$$

then, since (according to (2.44))

$$\begin{aligned} \int_I |y(t)|^2 dt &= \int_I |V_t(u)|^2 dt \leq \\ &\exp(\|u\|^2/\rho) \int_I \|V_t\|_{F_\rho(E)}^2 dt, \end{aligned} \quad (3.3)$$

it follows that $y \in L^2(I)$.

The following is easily established.

Theorem 3.1 [15]. $B_O^2(I, F_\rho(E))$ is a Hilbert space under the inner product

$$(V, W)_{B_O^2} \triangleq \int_I (V_t, W_t)_{F_\rho(E)} dt. \quad (3.4)$$

In connection with some applications, it is appropriate to introduce smoother versions of the space $B_O^2(I, F_\rho(E))$.

A function $g: I \rightarrow F_{\rho}(E)$ is said to be strongly differentiable at $t \in I$ if there is an element $Dg(t) \triangleq g'(t) \in F_{\rho}(E)$ such that

$$\lim_{h \rightarrow 0} \left\| \frac{g(t+h) - g(t)}{h} - g'(t) \right\| = 0. \quad (3.5)$$

Higher order strong derivatives $D^i g(t) \triangleq g^{(i)}(t)$, $i > 1$, are similarly defined.

Definition 3.2. $B_n^2(I, F_{\rho}(E))$ is the space of functions V from I to $F_{\rho}(E)$ such that $V, V', \dots, V^{(n-1)}$ are absolutely continuous and belong to $B_0^2(I, F_{\rho}(E))$, and $V^{(n)} \in B_0^2(I, F_{\rho}(E))$.

As in the case of Theorem 3.1, we have:

Theorem 3.2. $B_n^2(I, F_{\rho}(E))$ is a Hilbert space with the inner product

$$(V, W)_{B_n^2} = \sum_{i=0}^n a_i \int_I (V_t^{(i)}, W_t^{(i)})_{F_{\rho}(E)} dt, \quad (3.6)$$

where a_i are positive constants.

4. Nonparametric Nonlinear System Identification in the Noiseless Case:

Non-Causal Solution

We now return to the nonlinear system identification problem, and assume that, in the input-output relation

$$y = V(u) \quad (4.1)$$

satisfied by the system S to be identified, u and y belong respectively to E and $H_n^2(I)$, and $V \in B_n^2(I, F_p(E))$. For simplicity we will denote the spaces $H_n^2(I)$ and $B_n^2(I, F_p(E))$ simply by H_n^2 and B_n^2 .

One of the key points of our analysis is to reverse the roles usually assigned to u and V ; that is, we shall consider u to be an operator

$\tilde{u}: B_n^2 \rightarrow H_n^2$ which acts on V yielding y according to the relation

$$\tilde{u}(V) \triangleq V(u) = y. \quad (4.2)$$

Clearly, such an operator \tilde{u} is linear, and it is strongly continuous since, according to (2.44),

$$\|\tilde{u}(V)\|_{H_n^2} = \|V(u)\|_{H_n^2} \leq \exp(\|u\|_E^2/2p) \|V\|_{B_n^2}. \quad (4.3)$$

Also, we define the class C , introduced in Section 1, by

$$C = \{V \in B_n^2: \|V\|_{B_n^2} \leq \gamma\}, \quad (4.4)$$

where γ is a positive constant sufficiently large for C to have a nonempty intersection with the set

$$\chi = \{V \in B_n^2: \tilde{u}_i(V) = y_i, i = 1, \dots, m\}. \quad (4.5)$$

Condition (4.4) may be interpreted as a finiteness requirement on the gain of the system S (where we define the system gain as $\sup \{\|V(u)\|_{L^2(I)}^2 / \|u\|_E^2\} V(u)$) and on the smoothness of V , where the extent on the relative boundedness of the norms of the Frechet derivatives of V up to order n is determined by the constants a_i appearing in the definition of $\|V\|_{B_n^2}$ (see Theorem 3.2) and by v .

Let p be a functional on B_n^2 with the following properties.

- (i) p is bounded on bounded sets of B_n^2 ;
- (ii) $p \geq 0$ on B_n^2 ;

$$(iii) \quad p(V + W) \leq p(V) + p(W) \quad \forall V, W \in B_n^2;$$

$$(iv) \quad p(\alpha V) = |\alpha| p(V) \quad \forall \alpha \in \mathbb{R}, V \in B_n^2;$$

$$(v) \quad p \text{ is continuous on } B_n^2.$$

We define the error criterion e_C over the class C (introduced in Section 1) to be

$$e_C(V) = \sup_{W \in C} p(V - W), \quad (4.6)$$

and reformulate Problem 1 as:

Problem 1a. Same as Problem 1, with the additional specification that $Y = H_n^2(I)$ and that C and e_C be as defined by (4.4) and (4.6). ||

Remark. Equations (4.6) and (4.4) constitute a minmax criterion for the choice of the best \hat{V} . This type of criterion is particularly appealing when the number of measurements is small thus making a statistical criterion not plausible. ||

4.1. Geometrical Considerations

Our solution to Problem 1a relies on the geometry of the set

$$\Omega \stackrel{\Delta}{=} C \cap \chi, \quad (4.7)$$

to which we now turn our attention.

A set S in a normed linear space X is said to be symmetric if there exists an element $x_0 \in S$, called the center of S , with the property that $x_0 + \eta \in S \Rightarrow x_0 - \eta \in S$. Let p denote a seminorm on X with the properties (i) through (v) stated earlier (where now X replaces B_n^2) then the following result holds:

Lemma 4.1. If S is a bounded symmetric set in normed linear space X , then the center x_0 of S minimizes

$$e(x) = \sup_{y \in S} p(x - y). \quad (4.8)$$

Proof. Since p is a function of $(x-y)$ and S is symmetric, we may consider the translated set $S' = S - x_0$, which is also symmetric and has the center at the origin. We shall prove that the origin minimizes e on S' .

Let $\{x_i\}_{i=1}^{\infty}$ be a sequence in S' such that $p(x_i) \rightarrow e(0)$ as $i \rightarrow \infty$.

Then for any $\epsilon > 0$ there exists a $N(\epsilon)$ such that

$$e(0) \leq p(x_i) + \epsilon, \quad i \geq N(\epsilon) \quad (4.9)$$

and

$$\begin{aligned} 2p(x_i) &= p(2x_i) = p(x_i + y + x_i - y) \\ &\leq p(x_i + y) + p(x_i - y) = p(y - (-x_i)) + p(y - x_i). \end{aligned} \quad (4.10)$$

So we conclude that

$$\text{either} \quad p(y - x_i) \geq p(x_i), \quad (4.11a)$$

$$\text{or} \quad p(y - (-x_i)) \geq p(x_i). \quad (4.11b)$$

If (4.11a) is true, then

$$p(y - x_i) \geq p(x_i) \geq e(0) - \epsilon \quad (4.12a)$$

while if (4.11b) holds, we have

$$p(y - (-x_i)) \geq p(x_i) \geq e(0) - \epsilon. \quad (4.12b)$$

Now (4.12a) implies that

$$e(0) \leq p(y - x_i) + \epsilon \leq \sup_{x \in S'} p(y - x) + \epsilon \quad (4.13a)$$

and (4.12b) implies that

$$e(0) \leq p(y - (-x_i)) + \epsilon \leq \sup_{x \in S'} p(y - x) + \epsilon \quad (4.13b)$$

since x_i and $-x_i$ are elements of S' .

Because ϵ is arbitrary, it follows from (4.13a) and (4.13b) that

$$e(0) \leq e(y) \quad \text{for all } y \in S'. \quad \parallel$$

Now let

$$N = \{V \in B_n^2 \mid \tilde{u}_i(V) = 0, \quad i = 1, \dots, m\}. \quad (4.14)$$

The following constitutes a particularization of a well-known minimum norm result to our problem.

Lemma 4.2. If $V^* \in B_n^2$ satisfies the relations

$$(V^*, W)_{B_n^2} = 0 \quad \forall W \in N \quad (4.15)$$

and

$$\tilde{u}_i(V^*) = y_i, \quad i = 1, \dots, m, \quad (4.16)$$

then V^* is the unique solution of the minimization problem

$$\min_{V \in \chi} \|V\|_{B_n^2}^2, \quad (4.17)$$

(where χ is defined by (4.5)). \parallel

Next, we have

Lemma 4.3. Let V^* denote the solution of (4.17) and Ω be as defined in (4.7). If $V^* + \eta \in \Omega$ then

$$(V^*, \eta)_{B_n^2} = 0. \quad \parallel \quad (4.18)$$

Proof. Clearly, $V^* \in \Omega$.

To prove the lemma, suppose first that $\|V^*\|_{B_n^2}^2 = \gamma^2$. Then since V^* is a minimizer on Ω , it must be that

$$\|V^* + \eta\|_{B_n^2}^2 = \gamma^2, \quad (4.19)$$

which implies that

$$(V^*, \eta)_{B_n^2} = 0 \quad (4.20)$$

Suppose next that $\|V^*\|_{B_n^2}^2 < \gamma^2$. Then define the functional on \mathbb{R}

$$F(s) = (V^* + s\eta, V^* + s\eta)_{B_n^2}. \quad (4.21)$$

By continuity, there exists a neighborhood $[-\epsilon, \epsilon]$ about $s = 0$ such that

$$F(s) < \gamma^2.$$

Now since

$$\tilde{u}_i(V^* + \eta) = \gamma_i, \quad i = 1, \dots, m, \quad (4.22)$$

we have

$$\tilde{u}_i(\eta) = 0, \quad i = 1, \dots, m. \quad (4.23)$$

Hence $V^* + s\eta \in \chi$ for all s .

Since F is minimized at 0, we can differentiate (4.21) and obtain at $s = 0$,

$$F'(s)|_{s=0} = \{2(V^*, \eta) + 2s(\eta, \eta)\}_{s=0} = 0 \quad (4.24)$$

or

$$(V^*, \eta) = 0. \quad \parallel$$

Finally we have:

Lemma 4.4. If V^* and Ω are as in Lemma 4.3, then $V^* + \eta \in \Omega \Rightarrow V^* - \eta \in \Omega$ (and hence V^* is the center of Ω). \parallel

Proof. Since $V^* + \eta \in \Omega$, we have

$$\begin{aligned} \gamma^2 &\geq (V^* + \eta, V^* + \eta) = (V^*, V^*) + (\eta, \eta) \\ &= (V^* - \eta, V^* - \eta), \end{aligned} \quad (4.25)$$

the equalities following from the preceding lemma. \parallel

4.2. Solution and Algorithm

In order to state the system identification result that we are after, we require the following two additional lemmas.

Lemma 4.5. If $\{u_i: i=1, \dots, m\}$ is a set of distinct* elements of E , then $\{e^{u_i}: i=1, \dots, m\}$ is a linearly independent set of $F_\rho(E)$. \parallel

Proof. See Guichardet [11]. \parallel

Lemma 4.6. If $u_i, i=1, \dots, m$, are distinct elements of E , then the $m \times m$ matrix G with elements $G_{ij} = \exp \frac{(u_i, u_j)_E}{\rho}$ is nonsingular. \parallel

* u_1 and u_2 are distinct elements of E if $\|u_1 - u_2\|_E \neq 0$.

Proof. According to the preceding lemma, e^{u_i} , $i=1, \dots, m$, are linearly independent elements of $F_p(E)$. The result then follows from the fact

$$(e^{u_i}, e^{u_j})_{F_p(E)} = \exp \frac{(u_i, u_j)E}{\rho}. \quad \parallel$$

It is now possible to state:

Theorem 4.1. In Problem 1a, suppose that the probing inputs u_i , $i=1, \dots, m$, constitute a set of distinct elements of E . Then Problem 1a has a unique solution expressible in the form

$$\hat{V}(\cdot) = \sum_{i=1}^m c_i \exp \frac{(u_i, \cdot)E}{\rho}, \quad (4.26)$$

where $c_i \in H_n^2$ are determined by

$$c(t) \triangleq \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_m(t) \end{pmatrix} = G^{-1} \gamma(t) \triangleq G^{-1} \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \\ \vdots \\ \gamma_m(t) \end{pmatrix} \quad (4.27)$$

where G is the Gram matrix $\{\exp \frac{(u_i, u_j)E}{\rho}\}_{i,j=1, \dots, m}$ of the set $\{\exp((u_i, \cdot)/\rho) : i=1, \dots, m\}$.

Proof. Since u_i , $i=1, \dots, m$, are distinct elements of E , according to Lemma 4.6 G^{-1} in (4.21) exists.

Clearly, by construction,

$$\hat{V}(u_i) = \gamma_i, \quad i=1, \dots, m. \quad (4.28)$$

If $W \in N$, then

$$\begin{aligned} \langle \hat{V}, W \rangle_{B_n^2} &= \int_0^n \sum_{j=0}^n a_j D^j \sum_{i=1}^m c_i^*(t) \left(\exp \frac{u_i}{\rho}, D^j W_t \right) dt \\ &= \sum_{i=1}^m \int_0^n \sum_{j=0}^n a_j D^j c_i^*(t) D^j W_t(u_i) dt = 0 \end{aligned} \quad (4.29)$$

where the second and third equalities follow respectively from Theorem 2.3 and the fact that $D^j W_t(u_i) = 0$, $j=0, 1, \dots, n-1$, and $D^n W_t(u_i) = 0$ a.e. since $W \in N$. According to Lemmas 4.2 through 4.4, $\hat{V}(\cdot)$ defined by (4.26) is the center of Ω and hence by Lemma 4.1 the solution to Problem 1a. \parallel

Remark. It is of interest to obtain an estimate of the error

$$\xi = \|\hat{V}_t - V_t\|_{F_p(E)}^2, \quad (4.30)$$

where V_t and \hat{V}_t correspond respectively to the actual system and our estimate of it according to (4.26). According to the projection theorem, we have for ξ

$$\xi = (\|V_t\|_{F_p(E)}^2 - \tilde{Y}(t)^* G^{-1} \tilde{Y}(t)) \quad (4.31)$$

If $u_i, i = 1, \dots, m$, are orthonormal, then the diagonal elements of G are e (Napierian base) and its off-diagonal elements equal to unity. Then, (4.31) is expressible in the form

$$\xi = \|V_t\|_{F_p(E)}^2 - \alpha \sum_{i=1}^m |y_i(t)|^2 - \beta \sum_{\substack{i,j=1 \\ i \neq j}}^m y_i^*(t) y_j(t), \quad (4.32)$$

where

$$\alpha = \frac{e + m - 2}{e^2 + (m-2)e - (m-1)}, \quad \beta = \frac{-1}{e^2 + (m-2)e - (m-1)}. \quad (4.33)$$

An estimate of ξ can then be obtained by replacing γ^2 for $\|V_t\|_{F_p(E)}^2$ in the above formula, where γ is the constant introduced in connection with (4.4).

5. Causal Solution

Thus far, the only conditions assumed on the probing input-output pairs (u_i, y_i) , $i = 1, \dots, m$, are that they belong to appropriate spaces, and u_i , $i = 1, \dots, m$, constitute a set of distinct elements of E .

However, if we demand that the solution \hat{V} to the system identification problem of the preceding section be "causal", we need to impose additional restrictions on those pairs when $m > 1$; in other words, when the number of the probing input-output pairs is greater than one, the causality constraint on \hat{V} manifests as a set of admissibility restrictions on the pairs. These restrictions are developed in subsection 5.2. Also, in that subsection we introduce a weaker form of the concept of causality, which we call " ϵ -causality"; and show that it is possible to construct an ϵ -causal solution \hat{V} to the problem under consideration under less stringent admissibility requirements on the pairs (u_i, y_i) , $i = 1, \dots, m$, than for the strictly causal case.

For simplicity in presentation, we will assume that the operator V belongs to the space $B_0^2(I, F_p(L_2(I)))$ which we will abbreviate as B_0^2 , where $I = [0, 1]$, and hence (according to (2.46)) admits a representation of the form

$$y(t) = V_t(u) = h_0(t) + \frac{1}{1!} \int_0^1 h_1(t, s_1) u(s_1) ds_1 + \dots \\ + \frac{1}{n!} \int_0^1 \dots \int_0^1 h_n(t, s_1, \dots, s_n) u(s_1) \dots u(s_n) ds^{(n)} + \dots, \quad (5.1)$$

where $ds^{(n)} = ds_1 \dots ds_n$ and the h_i satisfy the conditions stated in connection with (2.46) and (2.47).

Let $P_t: L_2(I) \rightarrow L_2(I)$ be defined by

$$(P_t u)(s) = \begin{cases} u(s) & \text{if } s \leq t \\ 0 & \text{if } s > t \end{cases} \quad \text{a.e.} \quad (5.2)$$

Definition 5.1. $V \in B_0^2$ is causal if

$$P_t(V(P_t u)) = P_t(V(u)) \quad \forall t \in [0, 1]. \quad \parallel \quad (5.3)$$

Introduce the step function

$$w(t) = \begin{cases} 0, & t \leq 0 \\ 1, & t > 0 \end{cases}, \quad (5.4)$$

and define the operator $\tilde{C}: B_0^2 \rightarrow B_0^2$ by

$$\begin{aligned} (\tilde{C}V)(u)_t &= h_0(t) + \frac{1}{1!} \int_0^1 w(t-s_1) h_1(t, s_1) u(s_1) ds_1 \\ &+ \dots + \frac{1}{n!} \int_0^1 \dots \int_0^1 w(t-s_1) \dots w(t-s_n) h_n(t, s_1, \dots, s_n) \\ &\quad u(s_1) \dots u(s_n) ds^{(n)} + \dots \end{aligned} \quad (5.5)$$

Clearly, $\tilde{C}V$ is a causal operator. Furthermore it can be easily shown that \tilde{C} is a projection operator. In fact, it is obvious that $\tilde{C}^2 = \tilde{C}$, and to show that $\tilde{C}^{*+} = \tilde{C}$ simply use the definition of the inner product in B_0^2 .

Denote by M the subspace of causal operators in B_0^2 , that is

$$M = \tilde{C} B_0^2. \quad (5.6)$$

We seek the solution of;

Problem 1b. Same as Problem 1a, except that $E = Y = L_2(I)$, and hence $V \in B_0^2(I, F_\rho(L_2(I)))$ (abbreviated as B_0^2), and we require that V satisfy the constraint

$$V \in M. \quad \parallel \quad (5.7)$$

*Henceforth, the superscript $+$ on an operator symbol denotes its adjoint.

Let χ be as in (4.5) with B_n^2 replaced by B_0^2 . Clearly, $\chi \cap M$ is a closed linear variety in B_0^2 . Hence, if $\chi \cap M$ is nonempty, as we shall assume from now on, it follows from previous considerations that Problem 1b has a unique solution \hat{V} which is the minimum norm element of $\chi \cap M$.

To obtain an explicit representation for \hat{V} , we use the fact that \hat{V} is the unique element of $\chi \cap M$ orthogonal to $N \cap M$, where N is the subspace defined by (4.14) (with B_n^2 replaced by B_0^2). This fact is elicited by Lemma 5.1 below.

Let z denote an arbitrary element of $\chi \cap M$, i.e.

$$z \in \chi \cap M. \quad (5.8)$$

Clearly,

$$\chi = z + N. \quad (5.9)$$

Lemma 5.1. $\chi \cap M = z + (N \cap M)$.

Proof. Sufficiency: Let (using (5.9))

$$x \in \chi \cap M = (z + N) \cap M. \quad (5.10)$$

Hence, according to (5.8) and (5.9), x can be expressed as

$$x = z + r, \quad (5.11)$$

where $z \in M$ and $r \in N$. But since, according to (5.10), $x \in M$ and, as we have just stated, $z \in M$, it follows from (5.11) that $r \in M$. So the sufficiency is established.

Necessity: Let x be expressed as in (5.11). Then, according to the right side of Lemma 5.1, $x \in M$ because both z and r belong to M , and r belongs to N (because it is in $N \cap M$). Hence, by (5.10), $x \in \chi \cap M$. \parallel

5.1. The Case of a Single Probing Input-Output Pair

For the case of a single probing input-output pair, we have:

Theorem 5.1. If $m = 1$, the solution to Problem 1b is of the form

$$\hat{V}_t(u) = \tilde{\tau}_1(t) \exp\left(\frac{1}{\rho} \int_0^t u_1(s)u(s)ds\right), \quad (5.12)$$

where

$$\tilde{\tau}_1(t) = \left[\exp\left(-\frac{1}{\rho} \int_0^t |u_1(s)|^2 ds\right) \right] v_1(t). \quad (5.13)$$

Proof. It is clear that (5.12) is a feasible solution. According to Lemma 5.1, to show that it is of minimum norm it is sufficient to prove that such \hat{V} is orthogonal to NCM.

(5.12) can be expanded in the form

$$\begin{aligned} \hat{V}_t(u) = & \tilde{\tau}_1(t) \left(1 + \frac{1}{1!} \int_0^1 w(t-s_1) u_1(s_1) u(s_1) ds_1 + \right. \\ & \left. + \frac{1}{2!0^2} \int_0^1 \int_0^1 w(t-s_1) w(t-s_2) u_1(s_1) u_1(s_2) u(s_1) u(s_2) ds_1 ds_2 + \dots \right) \end{aligned} \quad (5.14)$$

Taking the inner product in B_0^2 of the above \hat{V} with an arbitrary V represented by (5.5) belonging to NCM, we obtain

$$\begin{aligned} (V, \hat{V}) = & \int_0^1 \tilde{\tau}_1(t) \left(h_0(t) + \frac{1}{1!} \int_0^1 h_1(t, s_1) w(t-s_1) u_1(s_1) ds_1 \right. \\ & \left. + \frac{1}{2!} \int_0^1 \int_0^1 h_2(t, s_1, s_2) w(t-s_1) w(t-s_2) \right. \\ & \left. u_1(s_1) u_1(s_2) ds_1 ds_2 + \dots \right) dt. \end{aligned} \quad (5.15)$$

In (5.15), because $V \in M$, we have

$$h_1(t, s_1) w(t-s_1) = h_1(t, s_1) \quad \text{a.e.} \quad (5.16a)$$

$$h_2(t, s_1, s_2) w(t-s_1) w(t-s_2) = h_2(t, s_1, s_2) \quad \text{a.e.} \quad (5.16b)$$

etc...

Using equations (5.16), expression (5.15) reduces to

$$\int_0^1 \tilde{c}_1(t) V_t(u_1) dt. \quad (5.17)$$

Above, $V_t(u_1) = 0$ a.e. because $V \in N$. Hence (5.17) vanishes thus establishing that \hat{V} defined by (5.12) is orthogonal to $N \cap M$. \parallel

5.2 The Case of Several Probing Input-Output Pairs

We now consider the case in which the number of the probing input-output pairs (u_i, y_i) , $i = 1, \dots, m$, is equal to or greater than two. We denote $\text{col}(u_1, \dots, u_m)$ and $\text{col}(y_1, \dots, y_m)$ respectively by \tilde{u} and \tilde{y} , and introduce the causal Gram matrix $\tilde{G}(t)$ with elements defined by

$$\tilde{G}_{ij}(t) = \exp\left\{\frac{1}{p} \int_0^t u_i(t') u_j(t') dt'\right\} \quad i, j=1, \dots, m. \quad (5.18)$$

5.2.1 Strictly Causal Solution

As in the case of (4.27) (with G now replaced by the causal $\tilde{G}(t)$), we require to obtain the solution $\tilde{c}(t) = \text{col}(\tilde{c}_1(t), \dots, \tilde{c}_m(t))$ of the equation

$$\tilde{G}(t) \tilde{c}(t) = \tilde{y}(t) \quad \text{a.e.} \quad (5.19)$$

The difficulty arises at the origin because $\tilde{G}(0)$ is a singular matrix with all elements equal to unity. According to this last fact, in order for (5.19) to have a solution at $t = 0$, $y_i(0)$, $i = 1, \dots, m$, must be well defined*, and there must hold

$$\tilde{c}_1(0) + \tilde{c}_2(0) + \dots + \tilde{c}_m(0) = y_1(0) = y_2(0) = \dots = y_m(0). \quad (5.20)$$

In addition, the following condition will be needed in the neighborhood of the origin:

$$\left. \begin{array}{l} \text{For some } \epsilon > 0, \text{ the restrictions of } u_i \text{ to } (0, t), 0 < t < \epsilon, \text{ are} \\ \text{distinct elements of } L^2(0, t) \text{ and} \\ \lim_{\substack{t \rightarrow 0 \\ t > 0}} G^{-1}(t) \tilde{y}(t) \text{ exists as a finite vector (where, in} \\ \text{taking the limit, we consider an appropriate member of the} \\ \text{equivalence class } \tilde{y}). \end{array} \right\} \quad (5.21)$$

We will call (5.20) and (5.21) a set of "admissibility conditions" on the probing input-output pairs.

Remark. For $m = 2$, condition (5.21) is obtained by requiring that

$$\left. \begin{array}{l} y_1'(0), y_2'(0), u_1(0), \text{ and } u_2(0) \text{ be well-defined} \\ \text{and finite and } u_1(0) \neq u_2(0). \end{array} \right\} \quad (5.22)$$

This is gleaned from the fact that (for $m = 2$)

$$\tilde{c}_1(t) = \frac{\tilde{G}_{22}(t) y_1(t) - \tilde{G}_{12}(t) y_2(t)}{\tilde{G}_{11}(t) \tilde{G}_{22}(t) - \tilde{G}_{12}^2(t)} \quad (5.23)$$

$$\tilde{c}_2(t) = \frac{\tilde{G}_{12}(t) y_1(t) - \tilde{G}_{11}(t) y_2(t)}{\tilde{G}_{11}(t) \tilde{G}_{22}(t) - \tilde{G}_{12}^2(t)}. \quad (5.24)$$

* If $y_i \in L^2(I)$, then y_i is an equivalence class. For the sake of complete generality, we say that $y_i(0)$ is "well defined" if a member \hat{y}_i of the class y_i is such that $\hat{y}_i(t)$ tends to a limit as $t \rightarrow 0$, and we denote this limit by $y_i(0)$. In other words, y_i is the equivalence class generated by such a \hat{y}_i . A similar statement applies when we say that a derivative of y_i is well defined at 0. $\tilde{c}(0)$ is defined as the solution of (5.19) with $\tilde{y}(0)$ defined as above.

Now, according to (5.18),

$$\lim_{t \rightarrow 0} \tilde{G}_{ij}(t) = 1, \quad i, j = 1, \dots, m \quad (5.25)$$

$$\lim_{t \rightarrow 0} \tilde{G}'_{ij}(t) = \frac{1}{p} u_i(0) u_j(0) \quad (5.26)$$

Hence, by l'Hôpital's rule,

$$\begin{aligned} \lim_{t \rightarrow 0} \tilde{C}_1(t) &= \lim_{t \rightarrow 0} \frac{\tilde{G}'_{22}(t) \gamma_1(t) + \tilde{G}_{22}(t) \gamma'_1(t) - \tilde{G}'_{12}(t) \gamma_2(t) - \tilde{G}_{12}(t) \gamma'_2(t)}{\tilde{G}'_{11}(t) \tilde{G}_{22}(t) + \tilde{G}_{11}(t) \tilde{G}'_{22}(t) - 2\tilde{G}_{12}(t) \tilde{G}'_{12}(t)} \\ &= \frac{\frac{1}{p} u_2(0) (u_2(0) - u_1(0)) \gamma_1(0) + \gamma'_1(0) - \gamma'_2(0)}{\frac{1}{p} (u_1(0) - u_2(0))^2} \end{aligned} \quad (5.27)$$

and similarly with $\tilde{C}_2(t)$, which shows that (5.22) is a sufficient condition for (5.21).

For the case of $m = 3$, the above type of calculation becomes extremely tedious. In application of l'Hôpital's rule, differentiation up to order three has to be carried out of the numerator and denominator of the expressions for $\tilde{C}_i(t)$, $i = 1, 2, 3$. Such a differentiation of the denominator gives

$$D'''(t) = \frac{d^3}{dt^3} \begin{vmatrix} \tilde{G}_{11}(t) & \tilde{G}_{12}(t) & \tilde{G}_{13}(t) \\ \tilde{G}_{21}(t) & \tilde{G}_{22}(t) & \tilde{G}_{23}(t) \\ \tilde{G}_{31}(t) & \tilde{G}_{32}(t) & \tilde{G}_{33}(t) \end{vmatrix}, \quad (5.28)$$

and hence

$$D'''(0) = \frac{3}{p^2} (u_2 - u_3) \begin{vmatrix} (2 u_1 u'_1 + \frac{1}{p} u_1^4) & u_1 & 1 \\ (u_1 u'_2 + u'_1 u_2 + \frac{1}{p} u_1^2 u_2^2) & u_2 & 1 \\ (u_1 u'_3 + u'_1 u_3 + \frac{1}{p} u_1^2 u_3^2) & u_3 & 1 \end{vmatrix}$$

$$\begin{aligned}
& + (u_1 - u_3) \begin{vmatrix} u_1 & (u_1 u_2' + u_1' u_2 + \frac{1}{\rho} u_1^2 u_2^2) & 1 \\ u_2 & (2 u_2 u_2' + \frac{1}{\rho} u_2^4) & 1 \\ u_3 & (u_2 u_3' + u_2' u_3 + \frac{1}{\rho} u_2^2 u_3^2) & 1 \end{vmatrix} \\
& + (u_1 - u_2) \begin{vmatrix} u_1 & 1 & (u_1 u_3' + u_1' u_3 + \frac{1}{\rho} u_1^2 u_3^2) \\ u_2 & 1 & (u_2 u_3' + u_2' u_3 + \frac{1}{\rho} u_2^2 u_3^2) \\ u_3 & 1 & (2 u_3 u_3' + \frac{1}{\rho} u_3^2) \end{vmatrix}, \quad (5.29)
\end{aligned}$$

where we have abbreviated $u_i(0)$ and $u_i'(0)$ by u_i and u_i' for $i = 1, 2, 3$.

From (5.29) it follows that $D^{(1)}(0) \neq 0$, if $u_1(0) \neq u_2(0) \neq u_3(0)$.

This and considerations resulting from differentiating the numerator lead to the following sufficient condition for (5.21) to hold:*

$$y_i'(0), u_i(0), \text{ and } u_i'(0), \quad i = 1, 2, 3, \quad \text{be} \quad (5.30)$$

well defined and finite and $u_i(0) \neq u_j(0) \forall i \neq j, i, j = 1, 2, 3$.

A sufficient condition similar to (5.30) can be derived for $m > 3$ but the calculation is too tedious to be carried out here. ||

Lemma 5.2. Suppose u_i and $y_i \in L^2(I)$, $i = 1, \dots, m$, satisfy conditions (5.20) and (5.21). Then (5.19) has a unique solution \tilde{c} , with $\tilde{c}_i \in L^2(I)$. ||

Proof. Since $\lim_{\substack{t \rightarrow 0 \\ t > 0}} \tilde{G}^{-1}(t) \tilde{y}(t)$ exists and is finite, there is a

subinterval $[0, \epsilon_1]$, $0 < \epsilon_1 < \epsilon$, over which $\tilde{c}(t) = \tilde{G}^{-1}(t) \tilde{y}(t)$ is bounded almost everywhere. Since u_i , $i = 1, \dots, m$, are distinct on $[0, \epsilon_1]$, they

* In this calculation, higher order derivatives (in ordinary or distributional sense) of u_i and y_i cancel out at 0.

are distinct on $[0, t]$, $\varepsilon_1 \leq t \leq 1$, and hence \tilde{G} and \tilde{G}^{-1} are continuous on $[\varepsilon_1, 1]$. This and the fact that $y_i \in L^2(I)$ establishes that the restriction of \tilde{z}_i to $[\varepsilon_1, 1]$ belongs to $L^2(\varepsilon_1, 1)$. Because, in addition, \tilde{z} is bounded a.e. on $[0, \varepsilon_1]$, we conclude that $\tilde{z}_i \in L^2(I)$. \parallel

The above leads to the following:

Theorem 5.2. Let $m \geq 2$ in Problem 1b. If, in this problem, we further restrict the operator to be identified \hat{V} to be such that (u_i, y_i) , $i = 1, \dots, m$, satisfy the admissibility conditions (5.20) and (5.21), then Problem 1b has a unique solution expressible in the form (with the understanding of the notation $\hat{V}(u)(t) = \hat{V}_t(u)$):

$$\hat{V}_t(\cdot) = \sum_{i=1}^m \tilde{z}_i(t) \exp\left[\frac{(u_i, \cdot)t}{\rho}\right], \quad (5.31)$$

where

$$(\cdot, \cdot)_t = (\cdot, \cdot)_{L^2(0, t)}, \quad (5.32)$$

and \tilde{z}_i are the components of the vector \tilde{z} obtained by solving (5.19).

Proof. It is clear that $\hat{V} \in B_0^2$ and

$$(\hat{V}u_i)(t) = y_i(t), \quad i = 1, \dots, m. \quad (5.33)$$

Proceeding in much the same way as in the proof of Theorem 5.1, we can show that each term of the form $\tilde{z}_i \exp\left[\frac{(u_i, \cdot)t}{\rho}\right]$ is orthogonal to every $V \in N \cap M$, this completing the proof. \parallel

5.2.2. ϵ -Causal Solution

In order to weaken the admissibility conditions (5.20) and (5.21), we introduce the following generalization of Definition 5.1:

Definition 5.2. $V \in B_0^2$ is ϵ -causal (for some positive $\epsilon < 1$) if

$$P_t(V(P_t u)) = P_t(V(u)) \quad \forall t > \epsilon \quad \parallel. \quad (5.34)$$

We may now formulate the following weaker form of Problem 1b:

Problem 1c. Same as Problem 1b except that the constraint (5.7) is replaced by the restriction that, for some specified ϵ , $0 < \epsilon < 1$, \hat{V} be non-causal for $0 \leq t \leq \epsilon$ and causal for $t > \epsilon$. \parallel

Theorem 5.3. Let ϵ , $0 < \epsilon < 1$, be such that the restrictions of u_i , $i = 1, \dots, m$, are distinct elements of $L^2(0, \epsilon)$. Then for this ϵ , Problem 1c has a unique solution described by

$$V_t(\cdot) = \sum_{i=1}^m c_i(t) \exp\left(-\frac{(u_i, \cdot) L^2(0, \epsilon)}{\rho}\right) \quad \forall 0 \leq t \leq \epsilon, \quad (5.35a)$$

$$V_t(\cdot) = \sum_{i=1}^m \tilde{c}_i(t) \exp\left(\frac{(u_i, \cdot)}{\rho} t\right) \quad \forall \epsilon < t \leq 1, \quad (5.35b)$$

where

$$c_i(t) = G^{-1} \tilde{\gamma}(t), \quad (5.36)$$

$$\tilde{c}_i(t) = \tilde{G}^{-1}(t) \tilde{\gamma}(t), \quad (5.37)$$

where G and $\tilde{G}(t)$ are $m \times m$ matrices with elements

$$G_{ij} = \exp\left(\frac{(u_i, u_j) L^2(0, \epsilon)}{\rho}\right) \quad (5.38)$$

$$\tilde{G}_{ij}(t) = \exp\left(\frac{(u_i, u_j)_{L^2(0, \epsilon)}}{\rho}\right) . \parallel \quad (5.39)$$

Proof. Clear from the arguments used in proving Theorems 4.1 and 5.2. \parallel

Remarks:

(a) In some cases, it may be easily verifiable that the condition on $\{u_i : i = 1, \dots, m\}$ of the above theorem holds for arbitrarily small ϵ , as in the example $\{u_i(t) = \sin kt, k = 1, \dots, m\}$.

(b) Also, it is immediately clear that a sufficient condition for the condition on $\{u_i\}$ of Theorem 5.3 to hold for arbitrarily small ϵ is that the u_i , $i = 1, \dots, m$, be continuous in the neighborhood of $t = 0$, and in addition $u_i(0) \neq u_j(0)$, $\forall i \neq j$.

6. Nonlinear System Identification in the Noisy Case

We next turn our attention to Problem 2 formulating it in the following specific form:

Problem 2a. Find the solution \hat{V} of

$$\min_{V \in B_n^2} \tilde{J}(V), \quad (6.1a)$$

where

$$\tilde{J}(V) = \|V\|_{B_n^2}^2 + \sum_{i=1}^m q_i^{-1} \|\tilde{u}_i(V) - y_i\|_Y^2. \quad (6.1b)$$

Remark: The class properties of V are taken into account by the constants a_i , $i=1, \dots, n$, (see equation (3.6)) picked in the definition of the inner product in B_n^2 , while the noise properties are accounted for by the weights q_i^{-1} , $i=1, \dots, m$, in (6.1b). ||

We seek a noncausal (not necessarily causal) solution \hat{V} of (6.1).

Details on the causal extension of the present results can be worked out in exactly the same way as in section 5 and hence will be omitted.

To obtain the solution to Problem 2a, we first construct a Hilbert space \tilde{B} defined by the Cartesian product

$$\tilde{B} = B_n^2 \times \underbrace{Y \times \dots \times Y}_{m \text{ times}}, \quad (6.2)$$

with the inner product between any two elements of \tilde{B} $f = \text{col}(f_0, f_1, \dots, f_m)$ and $g = \text{col}(g_0, g_1, \dots, g_m)$, where f_0 and $g_0 \in B_n^2$ and f_i and $g_i \in Y$, $i=1, \dots, m$, defined by

$$(f, g)_{\tilde{B}} = (f_0, g_0)_{B_n^2} + \sum_{i=1}^m q_i^{-1} (f_i, g_i)_Y. \quad (6.3)$$

Introduce the operator $L: B_n^2 \rightarrow \tilde{B}$ in the form

$$L = \text{col}(1, \tilde{u}_1(\cdot), \dots, \tilde{u}_m(\cdot)) \quad (6.4)$$

and denote by \tilde{y} the vector

$$\tilde{y} = \text{col}(0, y_1, \dots, y_m). \quad (6.5)$$

Then (6.1b) can be re-written as

$$\tilde{J}(V) = \|L V - \tilde{y}\|_{\tilde{B}}^2. \quad (6.6)$$

The following two lemmas, which result from elementary considerations, are stated without proof.

Lemma 6.1. The adjoints of $\tilde{u}_i: B_n^2 \rightarrow Y$ and $L: B_n^2 \rightarrow \tilde{B}$ are:

$$\tilde{u}_i^+ = \exp\left[\frac{(\tilde{u}_i, \cdot)}{\rho}\right] \quad (6.7)$$

and

$$L^+ = (1, q_1^{-1} \tilde{u}_1^+, \dots, q_m^{-1} \tilde{u}_m^+) . \quad (6.8)$$

Lemma 6.2. The minimizer \hat{V} of (6.6) satisfies the operator equation

$$L^+ L \hat{V} = L^+ \tilde{y} . \quad (6.9)$$

We also need the following result.

Lemma 6.3. The solution \hat{V} to Problem 1b belongs to N^\perp (orthogonal complement of N), where N is defined by (4.14).

Proof. Let P_0 denote the orthogonal projection from B_n^2 into N^\perp . Then

$$\begin{aligned} \tilde{J}(V) &= (P_0 V + (V - P_0 V), P_0 V + (V - P_0 V))_{B_n^2} \\ &\quad + \sum_{i=1}^m q_i^{-1} \|\tilde{u}_i(P_0 V) - y_i\|_Y^2 \\ &= (P_0 V, P_0 V)_{B_n^2} + \sum_{i=1}^m q_i^{-1} \|\tilde{u}_i(P_0 V) - y_i\|_Y^2 \\ &\quad + (V - P_0 V, V - P_0 V)_{B_n^2} \\ &= \tilde{J}(P_0 V) + \|V - P_0 V\|_{B_n^2}^2 \\ &\geq \tilde{J}(P_0 V) , \end{aligned} \quad (6.10)$$

with equality if and only if

$$V = P_0 V, \text{ i.e. if } V \in N^\perp. \quad \parallel$$

Now the set

$$\left\{ \exp\left[\frac{(u_i, \cdot)}{\rho} \right] : i = 1, \dots, m \right\} \quad (6.11)$$

spans N^\perp . Hence, by Lemma 6.2,

$$\hat{V}_t = \sum_{i=1}^m c_i(t) \exp\left[\frac{(u_i, \cdot)}{\rho} \right], \quad (6.12)$$

where $c_i \in Y$.

Substituting \hat{V} resulting from (6.12) into (6.9) and solving for c_i , $i = 1, \dots, m$, we are led to

Theorem 6.1. The solution to Problem 2a is

$$\hat{V} = c^T (I + G^+ Q^{-1} G)^{-1} G^+ Q^{-1} \tilde{y}, \quad (6.13)$$

where

$$c = \text{col}(c_1, \dots, c_m), \quad (6.14a)$$

$$Q = \text{diag}(q_1, \dots, q_m), \quad (6.14b)$$

$$\tilde{y} = \text{col}(y_1, \dots, y_m). \quad (6.14c)$$

7. Computer Simulation Example

In order to illustrate the preceding developments, we consider the simple example of a system described by the scalar differential equation:

$$\dot{y}(t) = -y(t) + u(t)y(t), \quad y(0) = 1, \quad (7.1)$$

on the interval $I = [0, 1]$. The exact input-output description for such a system is

$$\begin{aligned} y(t) &= V_t(u) = \exp \left\{ \int_0^t (u(s) - 1) ds \right\} \\ &= \exp(-t) \exp \left\{ \int_0^t u(s) ds \right\}. \end{aligned} \quad (7.2)$$

We assume we are given the eleven probing input-output pairs:

$$\begin{array}{ll} u_1(t) = 1 & y_1(t) = 1 \\ u_2(t) = \sin 2\pi t & y_2(t) = \exp(-t) \exp\{-(\cos 2\pi t - 1)/2\pi\} \\ u_3(t) = \sin 4\pi t & y_3(t) = \exp(-t) \exp\{-(\cos 4\pi t - 1)/2\pi\} \\ \vdots & \vdots \\ u_6(t) = \sin 10\pi t & y_6(t) = \exp(-t) \exp\{-(\cos 10\pi t - 1)/2\pi\} \\ u_7(t) = \cos 2\pi t & y_7(t) = \exp(-t) \exp\{(\sin 2\pi t)/2\pi\} \\ u_8(t) = \cos 4\pi t & y_8(t) = \exp(-t) \exp\{(\sin 4\pi t)/2\pi\} \\ \vdots & \vdots \\ u_{11}(t) = \cos 10\pi t & y_{11}(t) = \exp(-t) \exp\{(\sin 10\pi t)/2\pi\} \end{array} \quad (7.3)$$

With $p = 1$ and $u_i, \dots, i = 1, \dots, 11$ defined as above, (5.18) becomes

$$\begin{aligned} \tilde{g}_{ij}(t) &= \exp \left\{ \int_0^t u_i(s) u_j(s) ds \right\}, \\ i, j &= 1, \dots, 11. \end{aligned} \quad (7.4)$$

It is clear that the (u_i, y_i) , $i = 1, \dots, m$, satisfy the condition for an ϵ -causal solution \hat{V} to exist, given by the expression (5.35), for arbitrarily small ϵ (see Remark after Theorem 5.3).

Consider the test input $u(t) = t$. According to (7.2), the corresponding exact output is

$$y(t) = \exp(-t) \exp\left(\frac{1}{2} t^2\right) . \quad (7.5)$$

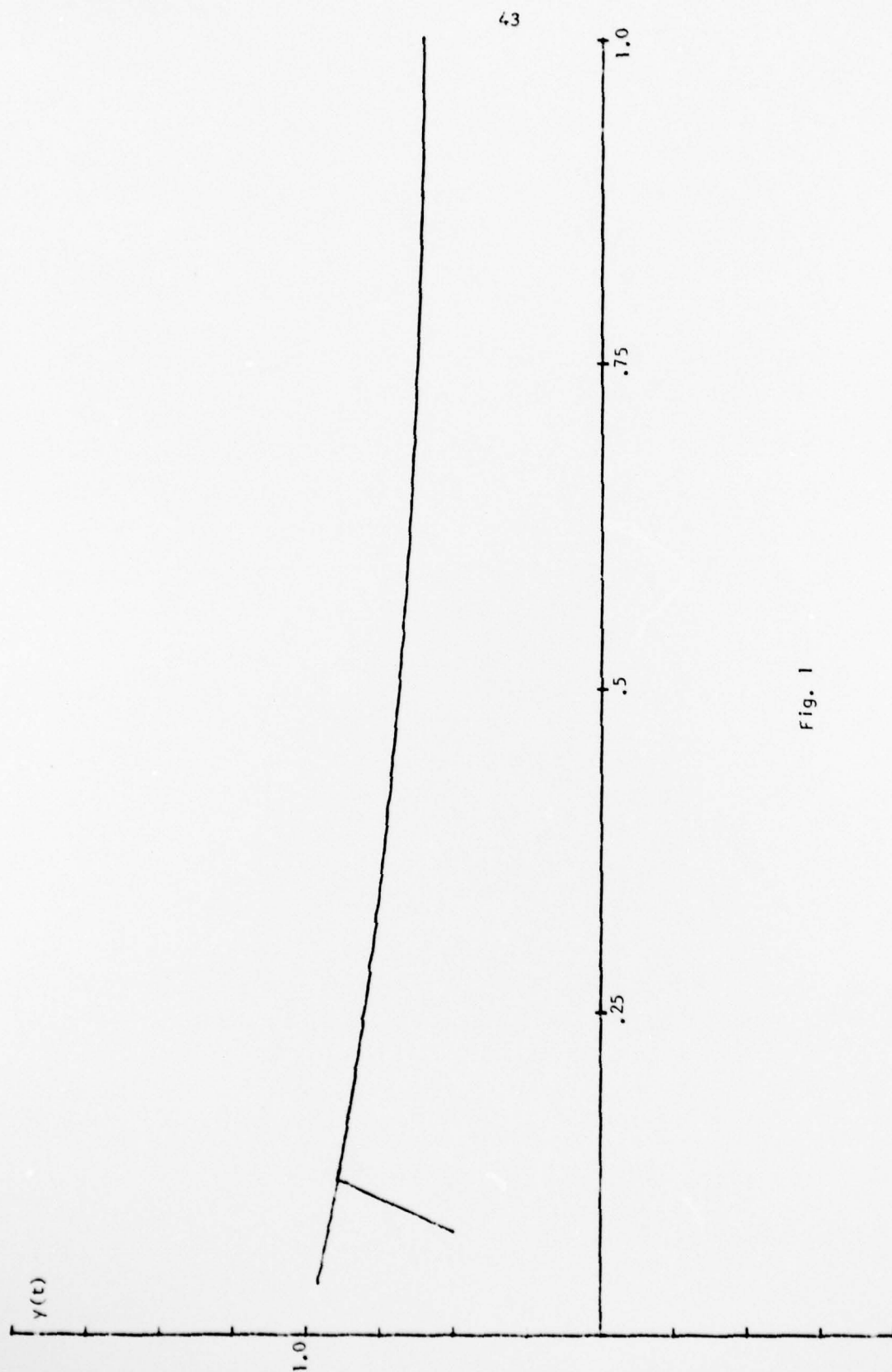
Table 1 and Fig. 1 compare samples of this output at sampling instants

$$t_i = (.04)i , i = 1, \dots, 25 , \quad (7.6)$$

with the corresponding samples of the output of our solution operator \hat{V} in a digital computer simulation, where we have assumed the value of $\epsilon = .04$. The agreement is remarkably good except for the first two instants. The disagreement near $t = 0$ results from the fact that $\det(\tilde{G}_{ij}(t))$ tends to zero as $t \rightarrow 0$. If we take $\epsilon = .12$, the agreement is nearly perfect.

TABLE 1

t_i	$y(t_i)$	$\hat{y}(t_i)$	Error $\left(\frac{y(t_i) - \hat{y}(t_i)}{y(t_i)}\right)\%$
.04	.96155838	2.80182753	-191
.08	.92607505	.50049397	45.95
.12	.89332931	.89325025	0.00885
.16	.86312134	.86312003	1.51×10^{-4}
.2	.83527021	.83527024	-3.59×10^{-6}
.24	.80961213	.80961212	1.23×10^{-6}
.28	.78599881	.78599882	-1.27×10^{-6}
.32	.7642961	.7642961	0
.36	.7443827	.7443827	0
.40	.72614904	.72614904	0
.44	.70949630	.70949630	0
.48	.69433550	.69433550	0
.52	.68058674	.68058674	0
.56	.66817845	.66817845	0
.60	.65704682	.65704682	0
.64	.64713523	.64713523	0
.68	.63839376	.63839376	0
.72	.63077882	.63077882	0
.76	.62425272	.62425272	0
.80	.61878339	.61878339	0
.84	.61434415	.61434415	0
.88	.61091344	.61091344	0
.92	.60847467	.60847467	0
.96	.60701608	.60701608	0
1.00	.60653066	.60653066	0



43

Fig. 1

8. Conclusion

In this paper we have developed an approach, based on a Fock space and Volterra expansion framework, for the formulation and solution of the nonlinear system identification problem, both without and under the causality and ϵ -causality constraints.

REFERENCES

- [1] Caprihan, A. and de Figueiredo, R.J.P., "The Generalized Smoothing Spline and System Identification," Internal Memorandum, Rice University, Dept. of Electrical Engineering, 1971.
- [2] de Figueiredo, R.J.P. and Caprihan, A., "An Algorithm for the Construction of the Generalized Smoothing Spline with Application to System Identification," Proc. of 1977 Conference on Information Sciences and Systems, The Johns Hopkins University, Baltimore, Md., pp. 494-500.
- [3] Balakrishnan, A.V., "Determination of Nonlinear Systems from Input-Output Data," Proc. Princeton Conference on Identification Problems in Communication and Control, pp. 31-49, March 1963.
- [4] Beutler, F.J. and Root, W.L., "The Operator Pseudoinverse in Control and System Identification" in Generalized Inverses and Applications (M.F. Nashed, Editor), Academic Press, New York, 1976.
- [5] Franklin, J.L., "Well Posed Stochastic Extensions of Ill-Posed Linear Problems," J. Math. Anal. Appl., 31, pp. 682-716, 1970.
- [6] Mosca, E., "On a Class of Ill-Posed Estimation Problems and a Related Gradient Iteration," IEEE Trans. on Automatic Control, AC-17, pp. 459-465, 1972.
- [7] Porter, W.A., "Data Interpolation, Causality Structure, and System Identification," J. Information and Control, vol. 29, no. 3, 1975.
- [8] Porter, W.A., "An Overview of Polynomial System Theory," IEEE Proceedings, Jan. 1976.
- [9] Porter, W.A., "Causal Realization from Input-Output Pairs," SIAM J. Control, vol. 15, no. 1, Jan. 1977.
- [10] Federer, H., Geometric Measure Theory, (Die Grundle. der Math. Wissensch., Bd. 15), Springer-Verlag, New York, 1969, page 43.
- [11] Guichardet, A., Symmetric Hilbert Spaces and Related Topics (Notes in Mathematics #261), Springer Verlag, New York, 1972.
- [12] Reed, M. and Simon, B., Methods of Modern Mathematical Physics (Functional Analysis, vol. 1, Academic Press, New York, 1973).
- [13] Dwyer, T.E., "Partial Differential Equations in Fischer-Fock Spaces for the Hilbert-Schmidt Holomorphy Type," Bull. of the Am. Math. Soc., vol. 77, pp. 725-730, 1971.
- [14] Dwyer, T.E., "Holomorphic Fock Representations and Partial Differential Equations on Countably Hilbert Spaces," Bull. of the Am. Soc., vol. 79, pp. 1045-1050, 1973.
- [15] Hille, E. and Phillips, R.S., Functional Analysis and Semi-groups, AMS Pub., Providence, 1957.